



SPI Japan 2016

# 開発者は不具合をどのように捉えているのか？

不具合票のテキストマイニングをはじめ

小島 義也

エプソンアヴァシス株式会社

e-mail:yoshiya.kojima@avasys.jp

**EPSON AVASYS CORPORATION**

# 目次

1. 背景と課題
2. 不具合に対する感情
3. ZUNDAとは？
4. ネガポジ分析・評価の実施
5. まとめと今後

# [背景] それぞれのギャップ

## ユーザーの不満足

- 使い勝手が悪い
- 処理状況がわかりにくい
- 新しいウィンドウが次々に開いてしまう

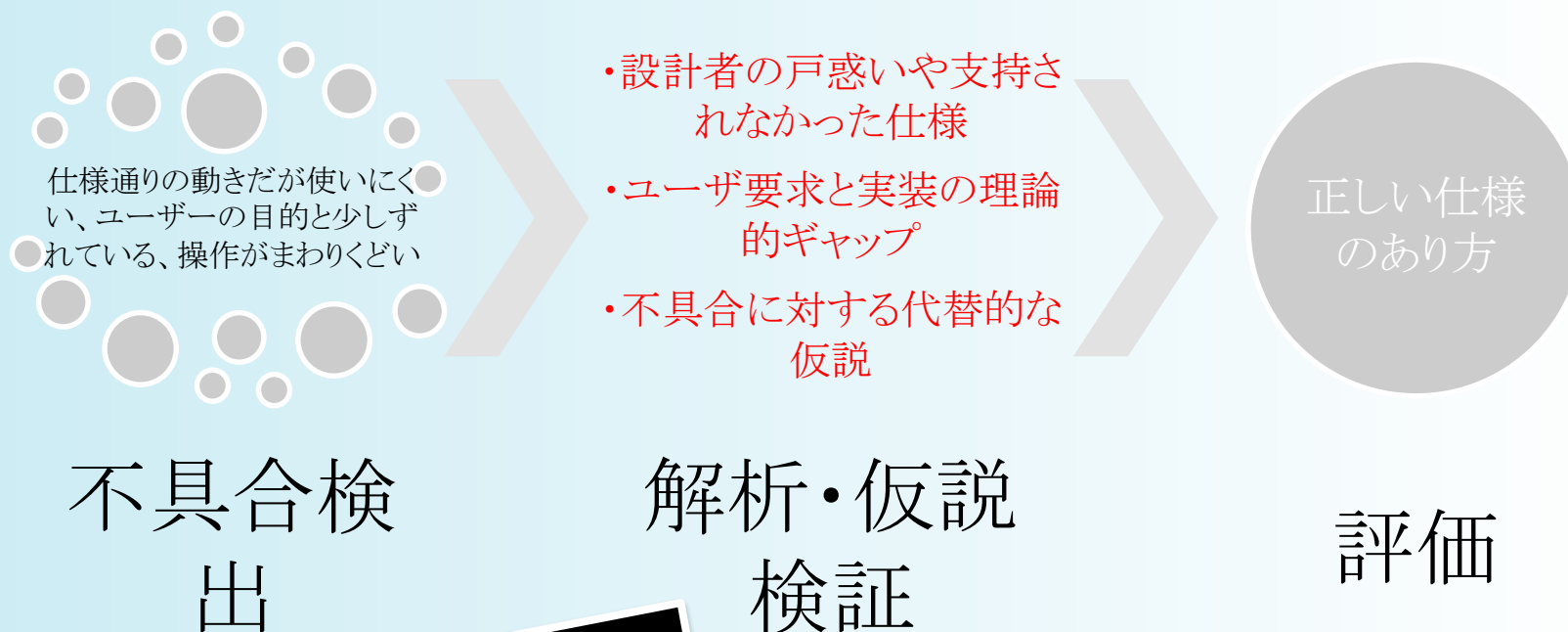
- 処理性能上げたのに
- 必要な情報は表示してあるのに
- わかりやすく誘導しているんだけど

## 開発者の不満足

仕様通りに作ったのに、できあがったものはユーザーが操作すると違って見えることが多々あります。

# 正しい仕様のあり方に至るプロセス

## ◆ ソフトウェア製品のβ版リリース後によく見られる不具合



製品が市場に出る前の品質を開発者がどのように捉えているのか？

# 開発者の不具合に対する「感情」

## 不具合票の項目

不具合票の本文・コメントには、肯定的な言葉（肯定的感情）や消極的な言葉（否定的感情）が見られる。

内容

トリアージ

障害度

定量化

ここが今までの分析の中心

優先度

検出機能

原因分類

備考

# 仮説と検証

⇒テキストマイニングによる  
ネガポジ評価を実施した。



①不具合票に記載された開発者の文章から、不具合に対する感情が捉えられないか？

②検出された不具合と感情の間には何か関連性はないか？

③その関連性には何か、開発プロセスの改善につながるヒントがあるのではないか？

不具合票にかならずある、フリー入力項目に注目。  
(例えば「不具合詳細」や「補足」「備考」といった文章情報)

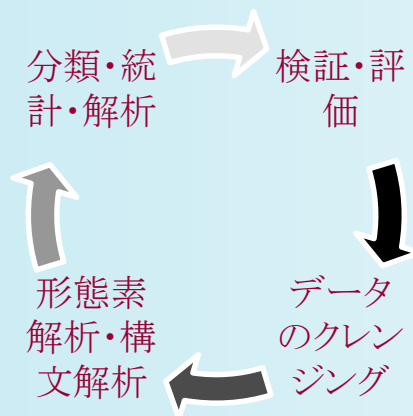
# テキストマイニング

## 概括

テキストマイニングとは、文字列から構成された大量のテキストデータを解析しその中から有益となる情報を取り出す技術のことです。今回は主に「自然言語処理」を行っています。

いまではテキストマイニング技術は、インターネット上に様々なツールが公開されていて、簡単にパソコン上で試すことができます。

実際には何をマイニングしたいかによって、手法やツールが異なっており、目的に応じて使い分けることが肝要です。



# ネガポジ評価

## 概要

ネガポジ評価は、Twitterの感情解析による株価推定の研究で注目を集め、有名になりました。[1]

ネガポジ評価は、自然言語処理の技術を利用して、テキスト内の文章を数値化し、ネガティブな感情を示す単語とポジティブな感情を示す単語等を数値化して算出しています。

一般にネガティブとポジティブの2種類に絞って分析を行うものをネガポジ評価と呼んでいます。

単語をネガポジを判別するためには、感情辞書を使用します。有名なのは高村大也先生の「スピンモデルによる単語の感情極性抽出」の研究成果として公開されている、「単語感情極性対応表」です。[2]

「単語感情極性対応表」はシンプルで使いやすい辞書ですが、もう少し細部な分析をしたいため、乾・岡崎研究室で開発された、ZUNDAというオープンソースで分析することにしました。[3]



# ZUNDA解析でのネガポジ判断の割合

		適用PJ 不具合 データ	Yahoo! 知恵袋	白書	新聞	書籍
価値判 断	0	94%	88%	84%	91%	89%
	ポジティ ブ	5%	10%	15%	8%	9%
	ネガティ ブ	1%	2%	1%	1%	2%

数字の"0"は、文章に特別な判断(モダリティ表現)が存在しない場合である。

不具合票は客観的な記述を要求されているため、新聞よりも高い数値となっている？

→理屈からいくと、"0"が高い方が感情的ではなく、不具合票としては優れているのか？

「拡張モダリティタグ付与コーパス作成の作業基準 version 0.8β」のデータを使用

# ZUNDA 日本語拡張モダリティとは？

**ZUNDA** 東北大学 乾・岡崎研究室で開発されたオープンソース日本語拡張モダリティ解析器です。文中のイベント(動詞や形容詞など)に対して、その真偽判断(イベントが起こったかどうか)、仮想性(仮定の話かどうか)などを解析します。

**モダリティ** 自然言語処理では、書き手が表明する態度や真偽判断、価値判断などの情報を、事象の拡張モダリティと呼んでいます。

詳しくは、以下の資料をご確認ください。

★Zunda: Japanese Extended Modality Analyzer

<<https://code.google.com/archive/p/zunda/>>

★「拡張モダリティタグ付与コーパス作成の作業基準 version 0.8β」

<<http://cl.naist.jp/nltools/modality/manual.pdf>>

# ZUNDA 拡張モダリティの項目

文章に存在する事象に対する、以下の 6 種類の項目をまとめて、拡張モダリティと呼ぶ。

項目	項目概要
態度表明者	対象とする事象の成否の判断や、他者への働きかけや問いかけをしている人物、もしくは、団体。多くの場合、この態度表明者は書き手である。
相対時	態度表明時から見た、対象事象の相対的な時間関係。過去・現在のことであるのか、それとも、未来のことであるのかを表す。
仮想	仮定された条件の有無。仮想世界の話であるのか、そうでないのかを表す。
態度	叙述、意志、働きかけ、問いかけなどの伝達的態度。言語学における「表現類型のモダリティ」のカテゴリーに相当する。
真偽判断	態度表明者による対象事象の真偽判断。対象事象が成立か不成立かを、確信度とともに表す。
価値判断	態度表明者による対象事象の価値判断。対象事象の成立が望ましいことであるかどうかを表す。

「拡張モダリティタグ付与コーパス作成の作業基準 version 0.8β」より引用

# ZUNDA 拡張モダリティのラベル

項目	ラベル一覧
態度表明者	筆者
相対時	未来、非未来
仮想	条件、帰結、0
態度	叙述、意志、欲求、働きかけ-直接、働きかけ-間接、働きかけ-勧誘、欲求、問いかけ
真偽判断	成立、不成立、不成立から成立、成立から不成立、高確率、低確率、低確率から高確率、高確率から低確率、0
価値判断	ポジティブ、ネガティブ、0

- ◆ 文章に記述されている情報は、単純な命題のみではなく、命題に対する情報発信者の主観的な態度も記述されている。
- ◆ ZUNDAは、解析した文に対し、上記の個々のラベルを設定している。
- ◆ ラベルの付与基準については、以下に詳しく説明されている。

「拡張モダリティタグ付与コーパス作成の作業基準 version 0.8β」より引用

# 分析・評価の流れ

1. ネガポジ評価の全体傾向
2. 仕様変更と不具合の比較
3. 不具合の欠陥起因分類(ODC)での比較
  - 人為的な原因に着目
4. 市場影響分類(ODC)での過去不具合との比較
  - 派生開発物としての傾向に着目
5. わかってきたこと

# ZUNDAでの不具合コメントへのラベル付与例

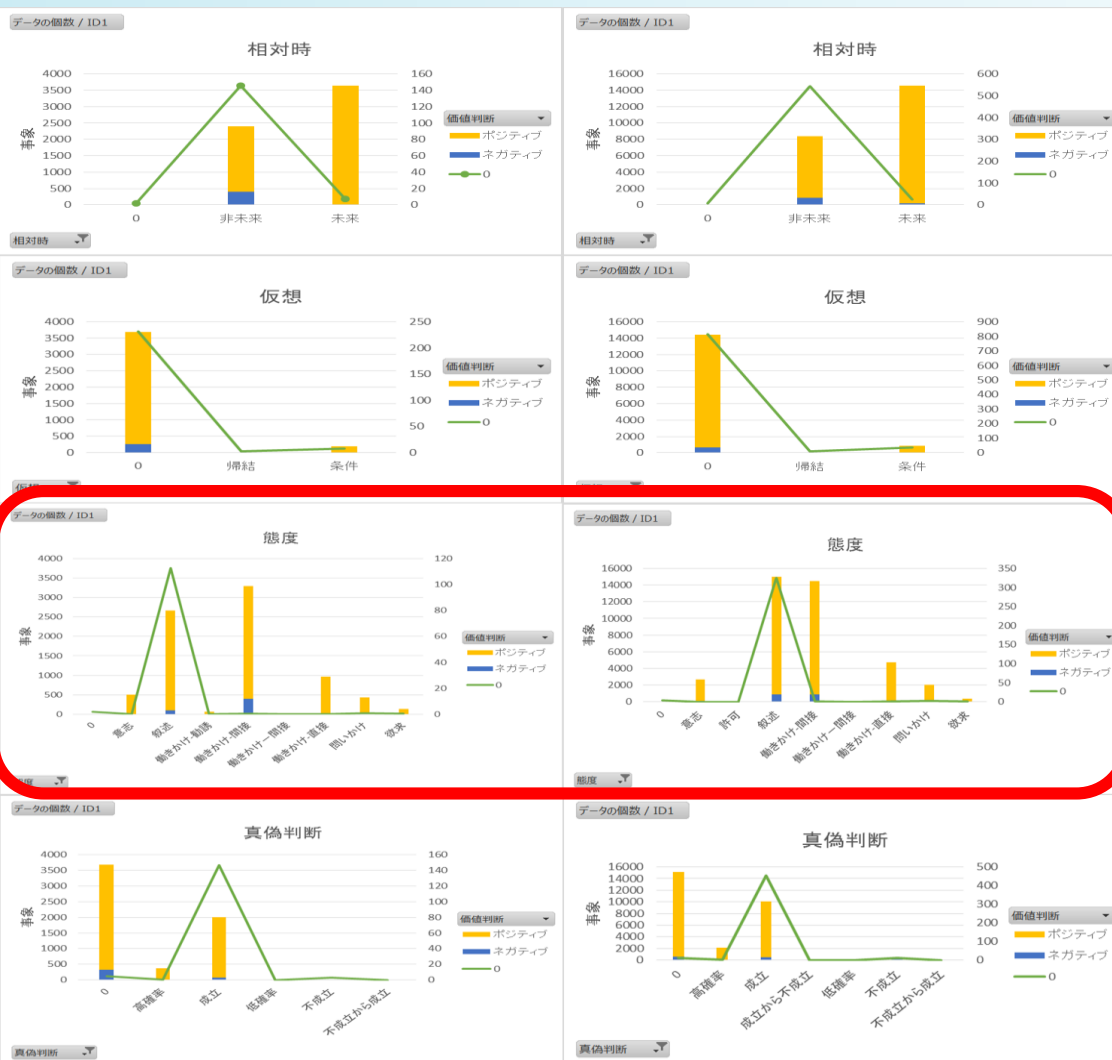
文

ユーザ設定で印刷できますので確かめてください。

態度表明者	相対時	仮想	態度	真偽判断	価値判断
Wr:筆者	未来	0	働きかけ- 直接	0	ポジティブ

- ◆ **ZUNDA**で不具合票の文を解析にかけると、上記のようなラベルが設定されてきます。
- ◆ 設定を数値に置き換えて、集計・分析をしています。

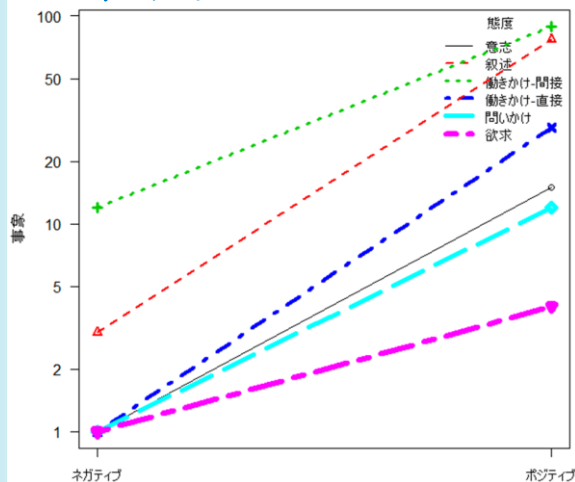
# 仕様変更と不具合の解析結果比較



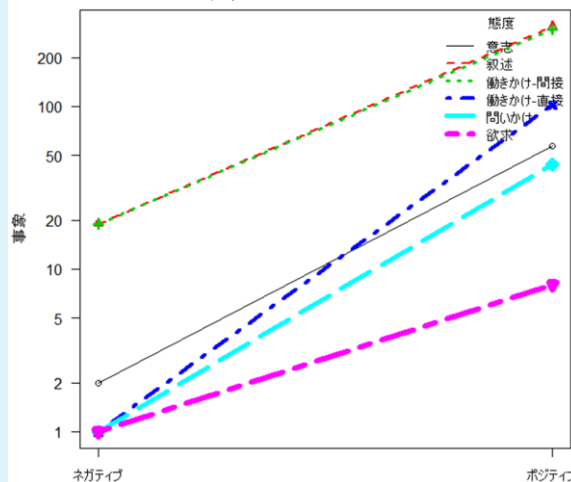
- ZUNDA解析結果で、仕様変更と不具合を比較して、どのモダリティが違うか？
- 「態度」の傾向が違ってくる。
- 次に、「態度」に焦点を当てて、傾向をみてみた。

# 仕様変更と不具合に対する開発者の「態度」比較

## 仕様変更



## 不具合

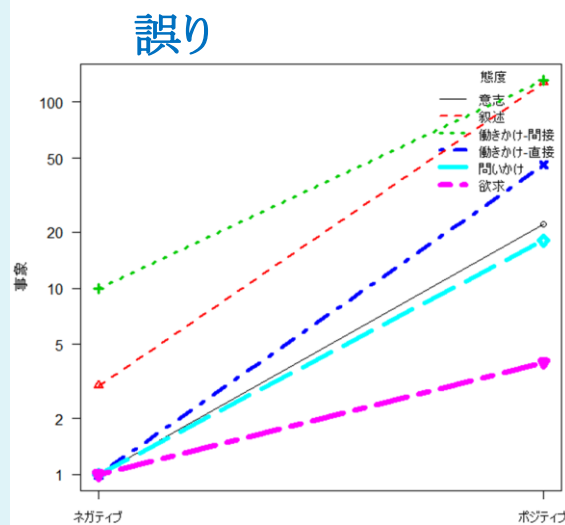
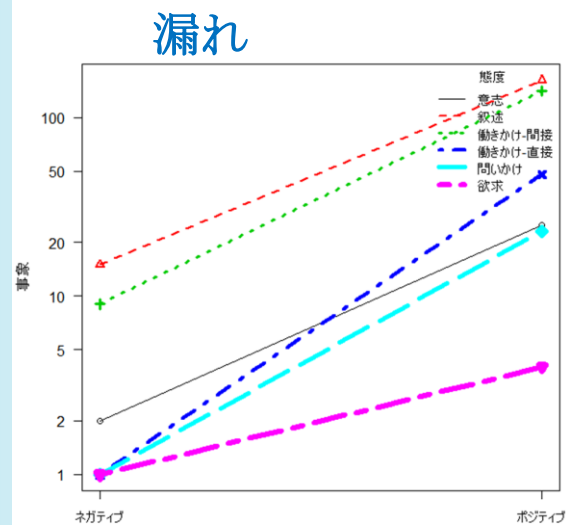


比較してみるとわかるのは、仕様変更は、「叙述」よりも「働きかけ-間接」が上回っている。

- 「叙述」は事象が表す内容や態度表明者の判断などを情報の受け手に伝えることを表す。文章に特別なモダリティ表現が存在しない場合である。
- 「働きかけ-間接」は態度表明者が、間接的な表現手段を用いて、行為の実行、もしくは、非実行を相手に求めることを表す。
- 行為の非実行を求めている場合には、「ネガティブ」表現になるため、仕様変更を望んでいない態度も記載されていることになる。
- 「叙述」より「働きかけ-間接」が上回っているのは、仕様変更に対する開発者の考えが多く記載されていることの現れと捉えられる。



# 不具合に対する開発者の「態度」



ODC分類のひとつに、欠陥起因という項目があり、これで不具合を起こした人為的な原因の「漏れ」と「誤り」に分けて、ネガポジ評価を行った。

ODCとは？  
SPIJapan2015参照[4]

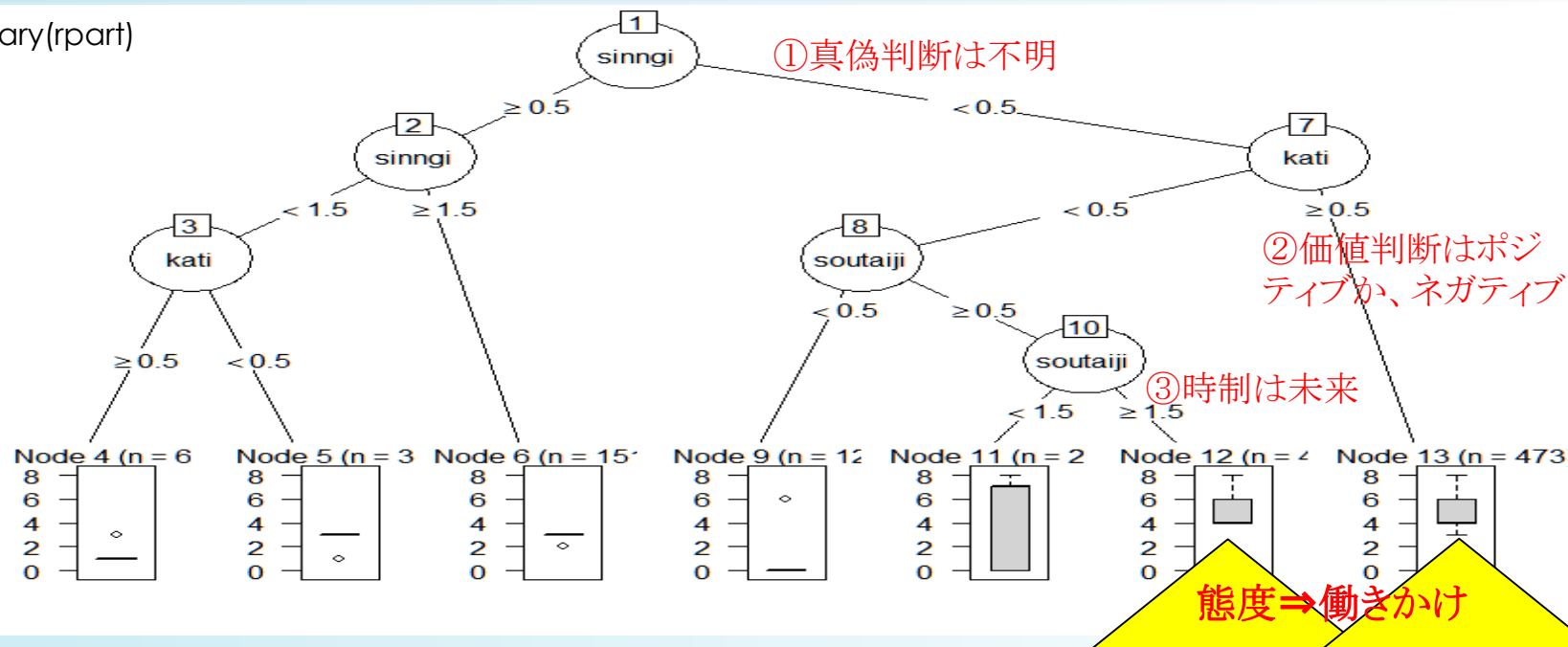
- 「漏れ」は、「叙述」が高くなっており、やはり設計漏れが原因の不具合であることを認識しての記載であったためと考えられる。
- またネガティブの「意志」も見られ、開発者が、ネガティブな自分自身の行為の実行／非実行を決定したことを表す事実もあった。
- ネガティブの「意志」は、漏れではなかったという記載もあり、実際には仕様としてのグレーゾーンであったことを示していたりする。
- 「誤り」は、仕様変更と同傾向であり、「働きかけ-間接」が高くなっており、ミスをした理由を記載しているためと考えられる。

# 「叙述」と「働きかけ－間接」

- 叙述よりも働きかけ－間接が上回っている場合(特にネガティブ)は、不具合に対する開発者の感情が働いており、問題性がある／あった不具合となる。
- 先に示した、「態度」では「働きかけ-間接」の高低で基準が示せそうではある。しかし、基準として使用するためにはもっと試行が必要ではある。
- 「働きかけ-間接」は、相手がいない状況において独り言としてつぶやける言語形式によって表現されることが多い、というモダリティであり、「漏れ」も「誤り」にしても、不具合を引き起こした理由を記載しているためと考えられる。
- そのため、「誤り」は不具合についての理由を述べる感情表現(ネガティブもポジティブも)が多かったと考えられる。

# わかってきたこと(働きかけのルール)

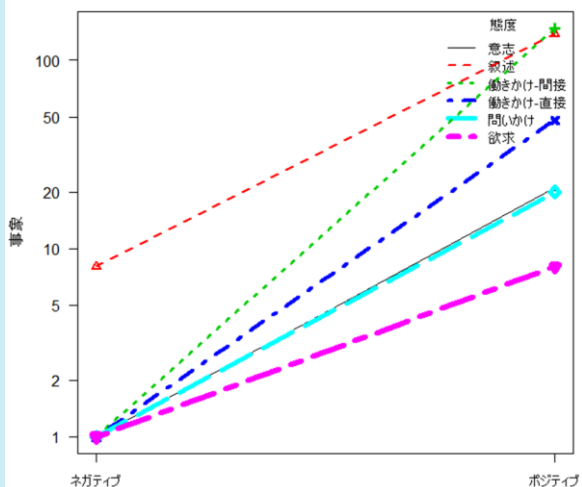
library(rpart)



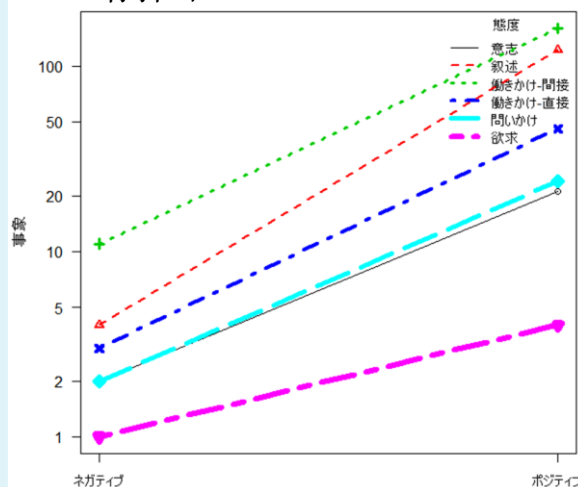
- ◆ ZUNDAで解析した不具合データを決定木にかけてみると、問題に対する態度の働きかけは、以下のルールで生じていることが多いことがわかる。
  - ① 真偽判断が詳細不明な文章である。
  - ② 価値判断はポジティブかネガティブのどちらかである。
  - ③ 価値判断がついていない場合で、時制は未来を語っているとき。
- ◆ 上記のルールから抽出される不具合票を中心に、現状のプロセス品質の変化点を推測することができないか？(課題)

# 市場影響分類での派生開発物の評価

今回version



前回version



「要件適合性」  
に当てはまる不  
具合を評価する

前はQCD未達。  
(事実)

- 今回と比較すると、前は「働きかけ-間接」のほかに、ネガティブな「働きかけ-直接」や「問いかけ」が目立っている。
- 例えば、グラフ比較から品質の傾向を予測することはできないか？あるいはネガティブとポジティブの差分から、改善の傾向が見えないか？
- 但し、前提としては、このデータでは開発者はあまり入れ替わっておらず、**不具合票の記載レベルは安定している**ことが条件のひとつと言える。

# ネガティブを軸に仮説をたてる

## ◆変化点の検出 — 不具合票のテキストに変化が起きたか否か

態度	要件適合性	信頼性	処理性能	ユーザビリティ	セキュリティ	移行性
意志	(+9)					
叙述	(+7)		---	---	---	---
働きかけ-間接	▲(+7)	▲				
働きかけ-直接	(+31)					
問いかかけ	(+10)					
欲求	(+3)					

(差分の差)

- 価値判断がネガティブ、態度が各項目が叙述よりも上回った場合に、▲をマークする。
- この▲が出てきたら、この市場影響(品質特性)に注意する。
- また、その隣に差分値(ポジティブ値－ネガティブ値)の差(今回－前回)を記載する。
- 程度問題として、今回が前回よりもプラスであれば、ポジティブ感が増えており、問題性は前回よりも低くなっていると考えられないか？

# 開発者は不具合をどのように捉えているのか？

## 特定の 態度

ある特定の不具合に対しては、開発者は特定の「態度」を持って対応していることがわかってきた。

うまく分析することができれば、開発者の感情から、プロジェクトのプロセス品質を推測することができそうであり、仮説・推測が正しければ、プロセス改善に使用することができる。

## 変化点 の検出

不具合票に対する開発者の感情の傾向(ふれ具合)は、推移を見たり、比較したりすることで、変化点を見つけることができる。ここから、プロジェクトの現況として捉えることができるのではないか、というのが今回のアプローチから見えてきたこと。

時系列での分析ができれば、例えば日報やレビュー議事録からも、プロジェクトの現況が読み取れるかもしれない？

# 妥当性確認と今後について

## 妥当性 確認

今回はプロジェクトの変化点検出の疑いありと判定する「ルール」を仮説として、人為的なミスなどを分析してみた。

このルールは他の不具合分類項目と組み合わせることで、分析に傾向を示すことができるのではないかと考えている。

例えば、「市場不具合」型や「欠陥原因」型など、各分類の特徴に応じたルールがあれば、プロジェクトの現況を推測できるようになるのではないかと考えている。

## 今後

感情評価のデータを継続したプロジェクトで評価結果を積み重ねることで、いくつかの観点から開発プロセスにおける変化点・特異点を見つけることができないか？

そうした感情評価のデータ傾向から、製品に対する新しい見方や評価を捉えることができ、ユーザビリティの向上につながる提案ができないか？



# 参考文献・資料

本稿の作成にあたり、以下の有用な知識やツールを公開頂いている方々に、この場をお借りして深謝申し上げます。

- ・ [1][ネガポジ解析によるWebデータと株価変動の相関関係評価]
  - 佐藤謙太;小高知宏;黒岩丈介;白井治彦
  - <<http://ci.nii.ac.jp/naid/120005588275>>
- ・ [2][スピンモデルによる単語の感情極性抽出]
  - 高村大也, 乾孝司, 奥村学
  - 情報処理学会論文誌ジャーナル, Vol.47 No.02 pp. 627--637, 2006.
- ・ [3]Zunda: Japanese Extended Modality Analyzer
  - Zundaは東北大学 乾・岡崎研究室で開発されたオープンソース 日本語拡張モダリティ解析器です。
  - <<https://code.google.com/archive/p/zunda/>>
- ・ [4][設計プロセスの課題をODC分析で改善してみた ～設計欠陥の原因追求による、欠陥修正工数の削減～]
  - 小島 義也 (エプソンアヴァシス株式会社)
  - <[http://www.jaspic.org/events/sj/spi\\_japan\\_2015/](http://www.jaspic.org/events/sj/spi_japan_2015/)>



御清聴ありがとうございました